Evaluation of Technology Concepts for Energy, Automation, and System State Awareness in Commercial Airline Flight Decks

Lynda J. Kramer[1], Timothy J. Etherington[2], Emory T. Evans[3], Taumi S. Daniels[4], Steven D. Young[5], and James R. Barnes[6]
NASA Langley Research Center, Hampton, Virginia, 23681, USA

Stefan R. Schuet[7] and Thomas Lombaerts[8]
NASA Ames Research Center, Moffett Field, California 94035, USA

and
Lance Sherry[9]
George Mason University, Fairfax, VA 22030, USA

A pilot-in-the-loop flight simulation study was conducted at NASA Langley Research Center to evaluate flight deck systems that (1) provide guidance for recovery from low energy states and stalls, (2) present the current state and expected future state of automated systems, and/or (3) show the state of flight-critical data systems in use by automated systems and primary flight instruments. The study was conducted using 13 commercial airline crews from multiple airlines, paired by airline to minimize procedural effects. Scenarios spanned a range of complex conditions and several emulated causal and contributing factors found in recent accidents involving loss of state awareness by pilots (e.g., energy state, automation state, and/or system state). Three new technology concepts were evaluated while used in concert with current state-of-the-art flight deck systems and indicators. The technologies include a stall recovery guidance algorithm and display concept, an enhanced airspeed control indicator that shows when automation is no longer actively controlling airspeed, and enhanced synoptic pages designed to work with simplified interactive electronic checklists. An additional synoptic was developed to provide the flight crew with information about the effects of loss of flight critical data. Data was collected via questionnaires administered at the completion of flight scenarios, audio/video recordings, flight data, head and eye tracking data, pilot control inputs, and researcher observations. This paper presents findings derived from the questionnaire responses and subjective data measures including workload, situation awareness, usability, and acceptability as well as analyses of two low-energy flight events that resulted in near-stall conditions.

[1] Aerospace Engineer, Crew Systems and Aviation Operations Branch, Mail Stop 152, Non-member.
[2] Engineering Technical Fellow, Rockwell Collins, Mail Stop 152, Senior Member.
[3] Aerospace Research Engineer, Flight Software Systems Branch, 5 North Dryden Street, Non-member.
[4] Aerospace Research Engineer, Electromagnetics and Sensors Branch, 8 North Dryden Street, Member.
[5] Aerospace Research Engineer, Safety-Critical Avionics System Branch, 1 South Wright Street, Fellow.
[6] Software Engineer, Booz Allen Hamilton, Mail Stop 159, Non-member.
[7] Computer Engineer, Intelligent Systems Division, Mail Stop 269-3, Member.
[8] Aerospace Engineer, Stinger Ghaffarian Technologies, Mail Stop 269-1, Associate Fellow.
[9] Director, Center for Air Transportation Systems Research, Mail Stop 4A6, Member.

I. Introduction

In the period 2010 to 2014, the Commercial Aviation Safety Team (CAST) sponsored a study of 18 commercial aviation events that occurred within ~10 years prior to the study kickoff. The results identified 12 recurring problem themes involving loss of airplane state awareness by the flight crew and suggested a number of intervention strategies [1]. CAST assessed these strategies for effectiveness and feasibility, and recommended Safety Enhancements (SEs) for the industry to implement [2]. Of these, six were deemed to require research to enable the enhancement. As part of its role in CAST, NASA initiated subprojects within the Airspace Operations and Safety Program to collaborate with the industry to produce many of these outputs.

The Automation and Information Management Experiments (AIME) are a series of experiments intended to address or achieve outputs defined within SE-207 [3] and SE-208 [4]. Together, the technologies under development and evaluation as part of the AIME series are intended to enable improved energy, automation, and/or system state awareness, as well as to provide new predictive capabilities with respect to these aspects of airplane state awareness (ASA).

As specified by CAST, the success criteria for all technology development research under SE-207 and SE-208 is to advance the technology readiness level (TRL) to five or greater. We interpret that as demonstration in a relevant environment and across a span of conditions such as those encountered in the events studied by CAST. In addition, any technology with a pilot interface should be judged usable and acceptable by pilots. We accomplish these goals using high-fidelity full-mission flight simulations. In these environments, not only can we replicate typical crew procedures and workload, but we can also replicate complex off-nominal situations such as those encountered during accidents. Such simulations also afford us the opportunity to test across a number of crews who may have disparate experience levels, expertise, and behavior tendencies.

The first AIME study (AIME-1) was completed in 2016, advancing and evaluating four technology concepts: a new synoptic display, predictive alerting of energy-related problems, automation mode change prediction and display, and dynamic maneuver envelope estimation and display [5-10]. AIME-2 is the subject of this paper. A complementary study (AIME-OKC) was completed at the FAA's Technical Center in Oklahoma City, OK, in September 2018. At least two more studies are underway or to be completed in 2019 (AIME-2.5 and AIME-3).

II. Test Overview and Objectives

AIME-2 was conducted over a 7-week period in April-May 2018. The objectives were four-fold. First, to raise the Technology Readiness Level (TRL) for new technology via testing in a high-fidelity flight simulator environment; this includes confirming that technologies perform as intended across a span of targeted conditions (e.g., representative of accident circumstances). Second, to evaluate the usability and acceptability of new technology concepts; thereby helping to decide whether the project is on the correct path, or needs a change of direction to produce the CAST desired research outputs. Third, to expose design characteristics that require refinement for future work; and discover previously unknown issues related to SE-207/208 goals. Fourth, to evaluate the use of the eye-tracking systems for their potential to help determine design effectiveness, as well as for detecting attention-related issues.

A. Technologies under Evaluation

One important factor to consider when evaluating new technology concepts is the reference platform into which the new technology will be inserted. Although retrofit on older aircraft may be possible, based on guidance from the CAST implementation team that retrofit would be expensive, it was decided that the Boeing 787 (B-787) flight deck [11] would be used for this purpose. As most of the relevant B-787 indicators, displays, and functions are also available on other recently developed aircraft, the findings reported here may be generalizable and applicable across these other platforms.

The B-787 is a highly advanced aircraft, representing in many ways, the state-of-the-art. The many indicators, displays, and functions relevant to SE-207/208 objectives are summarized in [5]. This list includes: the Primary Flight Display (PFD), the Head-Up Display (HUD), the Flight Mode Annunciator (FMA) in the upper center of the PFD, airspeed and altitude "tapes" on the left and right side of the PFD, the Flight Path Vector (FPV) on the PFD, the Navigation Display (ND), the Vertical Situation Display (VSD), the Engine Indicating and Crew Alerting System (EICAS), the synoptic displays (7 types) available on Multi-Function Displays (MFDs), the Lower MFD (LMFD) which serves as the pilot interface to the Flight Management System (FMS), Electronic Checklists (ECL), the Electronic Flight Bag (EFB), the Air Traffic Control (ATC) display that lists ATC data link messages that have been received. The surfaces associated with these displays were organized in the flight simulator to mimic the B-787 layout as shown in Fig. 1.

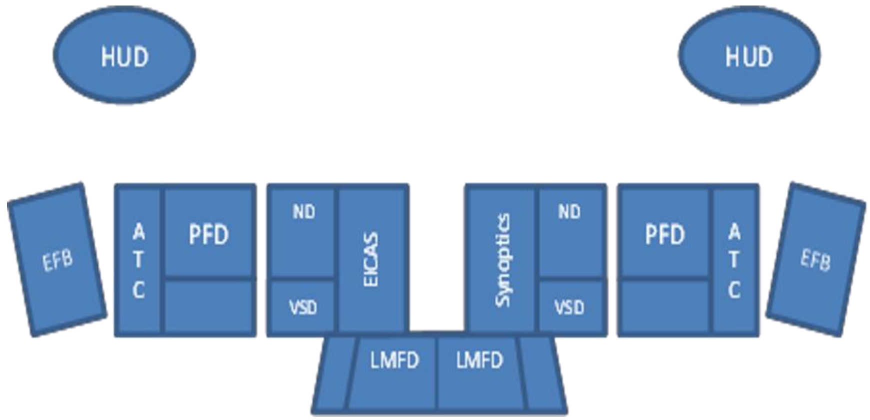

Fig. 1 Display surface reference layout

As extensions to this reference set of displays and functions, three new technology concepts were evaluated. These are referred to as: (1) Enhanced Synoptic (eSyn) pages with associated Simplified Electronic Checklists (sECL). The sECLs are shortened versions of standard ECLs to be used in conjunction with the eSyn pages without any loss of information, (2) EACI - Enhanced Airspeed Control Indicator, a visual indication that the automation is no longer actively controlling to an airspeed target, and (3) SRG - Stall Recovery Guidance, visual cues on the PFD indicating pitch and roll control inputs as well as throttle inputs to recover from stall or near-stall situations.

SRG and eSyn/sECL are the result of previous research [12-16] and have matured to a relatively high TRL. For these, AIME-2 sought to evaluate their span of applicability across a broader set of conditions, such as those encountered in the CAST events. EACI had a lower TRL [17] with AIME-2 being the first opportunity for it to be demonstrated in full-mission conditions and assessed for usability/acceptability in such conditions. Below, these three technologies are briefly explained. Additional descriptions of their designs for this study and findings resulting from their use can be found in [18-20].

1. Enhanced Synoptics and Simplified Electronic Checklists

Two types of eSyn pages (Fig. 2) were evaluated in this study: 1) enhancements to existing synoptic pages [12] and, 2) new synoptic pages for failures in which no relevant synoptic exists [12, 13]. B-787-like synoptic displays were enhanced to depict and provide additional information regarding failures and effects. These enhancements enabled the simplification of associated ECLs by removing information now provided on the enhanced synoptic display as well as providing context relevant data on the checklists.

For example, in this study, the Flight Control Synoptic page (Fig. 2, left image) was enhanced to illustrate the effects of a left hydraulics system failure and the associated Hydraulic Systems Pressure (Left Only) non-normal checklist was reduced in length (by removing text information now available on the synoptic). The Flight Control eSyn page pictorially showed the effects (left autopilot and left thrust reverser inoperative, reduction in ailerons, manual speed brakes, alternate flaps extension, alternate gear extension, etc.) of the failed left hydraulic system thus enabling a 35% reduction in the ECL length since many checklist items/notes were now shown on the Flight Control eSyn page.

For some types of failures, there is no relevant existing synoptic. In these cases, a new synoptic page could be defined. The example used for this research was loss of flight critical data (e.g., airspeed) provided by the air data system and inertial reference units. A new eSyn page, referred to here as the System Interaction Synoptic (SIS), was created that graphically depicted whether flight-critical data was valid, the operational state of the sources of the data, and the effects on systems that received the data (Fig. 2, right image).

Unique eSyn pages are associated with specific aircraft faults indicated by an EICAS message. The eSyn color scheme matches that of the other synoptic pages. Data or data paths shown in green indicate valid data, those shown in white indicate operable data but with reduced quality, and those in amber indicate invalid data. The eSyn is available to the flight crew as a quick-look reference regarding failures and their effects throughout the flight.

For example, during AIME-2 an EICAS message "Airspeed Unreliable" resulted in a change of information on the eSyn page (SIS synoptic shown on the right-hand side in Fig. 2) indicating bad data from both air data computers (ADCs) due to a blocked pitot-static system. The impact of this fault is shown on the eSyn as the auto-flight system is inoperable (amber text in symbolic MCP near top of SIS), alternate altitude and airspeed is derived from GPS and angle of attack, respectively, (white text in the flight critical data type box in the center of the SIS), and the aircraft has reverted to secondary flight control mode (amber text at bottom of SIS) requiring manual flight. Additionally, pitch attitude and power settings based on the current aircraft configuration are shown on the SIS page instead of requiring the flight crew to spend time locating and interpolating the data in the appropriate table of their Quick Reference Handbook.

Associated with the EICAS notification and the eSyn page is a sECL. This shortened checklist is intended to be used with the eSyn to enable the pilot monitoring (PM) to complete the non-normal ECL in less time and with greater understanding of failure effects.

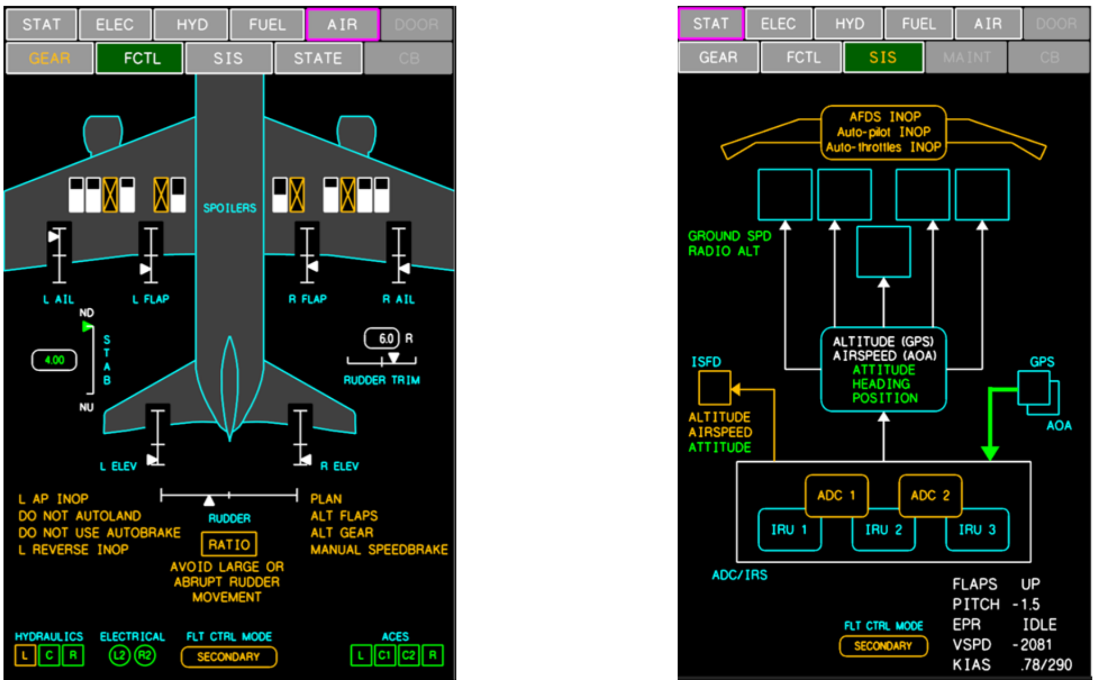

Fig. 2 Enhanced flight control synoptic for left hydraulic systems failure (left image) and new synoptic for loss of air data failure and effects (right image)

2. Enhanced Airspeed Control Indicator

The EACI technology concept is an enhancement to the PFD's airspeed tape (Fig. 3) to indicate whether the auto-flight system (including auto-throttles) is actively controlling to a target airspeed. Previous research [17] suggested new indicators near the airspeed tape as well as the use of color to indicate which, if any, system was controlling to a target airspeed. For AIME-2, the concept was simplified to only indicate when airspeed was not being actively controlled by the automation,

1. White XX's were shown at the commanded speed indicator at the top of the airspeed tape
2. White XX's were shown at the boxed current airspeed indicator in the center of the airspeed tape
3. White XX's were shown at the airspeed bug location as set in the MCP or FMS

3. Stall Recovery Guidance

SRG [15, 16] provides flight director-like guidance to aid pilots to escape from approach-to-stall and stall situations (Fig. 4). Beginning at stick shaker onset, the following guidance cues and indicators were provided on the PFD:

1. Red 'SRG' indicated across the top replacing all three modes in the FMA
2. Red 'RECOVER' indicated across the bottom of the PFD
3. Magenta dual-cue flight director for pitch and roll recovery

4. White and magenta throttle guidance cues near the right wing symbol (indicating whether to advance or retard the throttles)

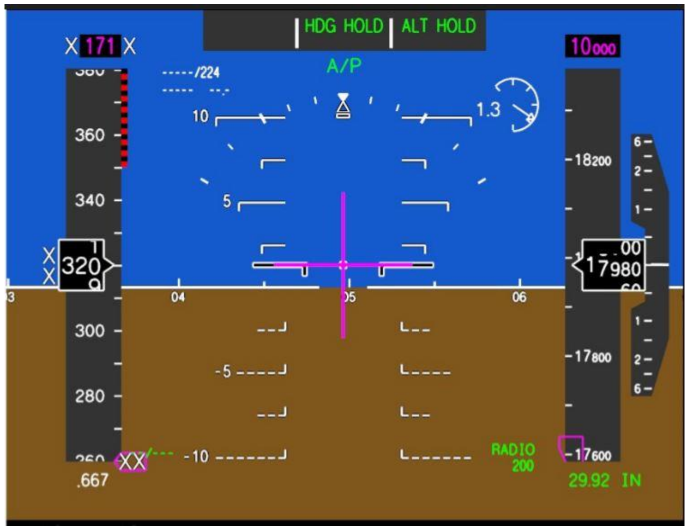

Fig. 3 Enhanced airspeed control indicator (as implemented on the PFD for AIME-2)

For AIME-2, reset of the PFD to normal mode after recovery was pilot-initiated through any MCP selection. Further, two different SRG algorithms were evaluated – an energy-based approach and a model-predictive control approach. The energy-based approach relied directly on the flight physics of the aircraft and was independent of aircraft model information [15]. For the model-predictive control approach, an optimal recovery was computed using a fast convex optimization algorithm [16]. Detailed descriptions of these algorithms can be found in [19, 20]. The results reported here do not differentiate between the two algorithms. Additionally, the SRG algorithm and display concept was active for all data collection runs in case the pilots experienced a stick shaker event while evaluating the EACI or eSyn/sECL display concepts. It provided a safety net from which pilots could recover from a low-energy situation.

Fig. 4 Stall recovery guidance display concept (as tested in AIME-2)

B. Flight Simulation Facility and Environment

AIME-2 utilized the Research Flight Deck (RFD) simulator (Fig. 5) within the Cockpit Motion Facility at NASA's Langley Research Center. Although initially designed as a B-757 flight deck, the RFD was modified to create the CAST-recommended reference test condition (i.e., B-787-like displays, interfaces, and functions). The simulator is configured with four 10.5-inch Vertical (V) by 13.25-inch Horizontal (H), 1280x1024 pixel resolution color displays, tiled across the instrument panel. The RFD also includes dual Rockwell Collins HGS-6700 HUDs, MCP, FMS, and EFBs for Pilot Flying and Pilot Monitoring. Two five-camera Smart Eye™ head and eye tracking systems are installed to quantify both crew member's head movement and eye-gaze behavior. Both eye tracking systems data outputs and the simulator state data output are time-synchronized.

The full-mission RFD utilizes a Boeing B-757-200 aircraft model, albeit controlled through sidestick inceptors. The pilot and co-pilot inceptors are directly linked as if mechanically connected. The auto-throttle system backdrives the throttle handles to directly reflect the power setting commanded to the engines. Take-off, go-around (TOGA) buttons and auto throttle disconnect buttons are placed on the throttle handles. A collimated out-the-window (OTW) scene is produced by a Rockwell Collins Image Generator (IG) graphics system providing approximately 200 degrees (deg) H by 40 deg V field-of-view at 26 pixels per degree.

Several operational complexities were also enabled for AIME-2. Flight crews were immersed in high density traffic and adverse weather environments that included many concepts either currently emerging in the industry or planned for the near future (e.g. digital data link services, synthetic vision systems, and area navigation/required navigation performance operations). In addition, the study emulated off-nominal (and complex) situations such as unexpected weather events, traffic deviations, equipment failures, unexpected clearances, and changes to flight plans. This provided a realistic operational environment, albeit a complex one, in which flight crews may have little prior experience or training. However, previous similar experiments have shown crews can learn and perform well with limited exposure to this environment [21].

Fig. 5 RFD simulator

C. Scenarios

Given the test objectives, one of the most challenging aspects of the AIME-2 setup was defining a relevant set of flight scenarios for the crews to experience. Advancing TRL requires testing across a good range of conditions, particularly any conditions that expose the need for the new technologies being developed. Similarly, effective usability assessment by the crew can only be done if the evaluation conditions span situations where usage would occur. Actual events in the industry are rare and often only happen when multiple (typically up to seven or more) things go wrong. Reproduction in the simulation environment requires a complex interaction of event, distraction, and sometimes suppression of current alerting. With this in mind, several sources were used to develop scenarios for the testing. Specifically, they were:

- CAST ASA study events [1][2]
- Accident reports from recent events [22], [23], [24]
- Previous research on eSyn/sECL [12,13], SRG [15, 16], and EACI [17]
- Previous complexity and automation studies [13, 21]
- Interviews and workshops with industry Subject Matter Experts (e.g., pilots and designers)

Based on these sources, four types of scenarios were defined for AIME-2: nominal (type 1); system failures (type 2); increased risk of low energy (type 3); and distractions affecting workload management (type 4). These scenario types were not designed to be mutually exclusive (e.g., auto throttles failure (type 2) coupled with an ATC-issued runway change (type 4) may, and did, induce an increased risk of low energy (type 3)). Following multiple workshops and SME consultations, nine scenarios were developed using these scenario types across a span of weather conditions. In terms of the test matrix and technology evaluations, each crew flew one SRG training scenario (three times by each pilot), two eSyn/sECL scenarios, three EACI scenarios and three SRG scenarios. These scenarios are described below.

1. Enhanced Synoptics and Simplified Electronic Checklists Scenarios

Two off-nominal runs were flown to evaluate the eSyn/sECL concept. The off-nominal runs involved either a left hydraulic systems failure or unreliable airspeed information due to a blocked pitot-static system. The runs ended once the non-normal checklists associated with the failures were completed.

<u>Hydraulic Failure</u>

The scenario was initialized at 18,000 ft Mean Sea Level (MSL) in cruise configuration (Landing Gear UP, Flaps UP) as the crew prepared for their initial descent for the RNAV RNP 31R approach at JFK. Seven minutes after run start, a hydraulic leak occurred that prevented the left engine and left electric hydraulic pumps to supply hydraulic pressure. In addition, the failure prevented the power transfer unit from providing system pressure as well. The left hydraulics system failure resulted in the following items inoperative: left autopilot, left thrust reverser, normal flap extension, and normal gear extension. The crew was required to complete the Hydraulic System Pressure (Left only) non-normal checklist and to perform alternate flap extension and alternate gear extension.

<u>Pitot-Static Failure</u>

The scenario was initialized at 18,000 ft MSL in cruise configuration (Landing Gear UP, Flaps UP) as the crew prepared for their initial descent for the RNAV RNP 13L approach at JFK. At run startup, an EICAS message “HEAT PITOT L-C-R” was triggered for failure of the pitot heat system. The aircraft was in weather with precipitation at that time. At 10 mi from waypoint CAMRN, an ATC datalink message “HOLD AT CAMRN AS PUBLISHED, 2 MINUTE LEGS, MAINTAIN 210 KNOTS, EXPECT FURTHER CLEARANCE IN 30 MINUTES, DESCEND IN THE HOLD TO 6000” was issued. Descending through 9,000 ft MSL, a total failure of the pitot-static system occurred due to icing/blockage of the pitot and static ports as a result of the annunciated failure of the pitot heat system. This condition triggered an EICAS message “UNRELIABLE AIRSPEED”, as airspeed and altitude information became unreliable and their sources switched to angle-of-attack (AOA) airspeed and global positioning system (GPS) altitude, respectively, for both Captain and First Officer PFDs. This failure also caused the autopilot, auto throttles, and flight directors to become inoperative. The crews had to manually fly the aircraft (without flight guidance or auto throttles) while they completed the six page Unreliable Airspeed ECL.

2. Enhanced Airspeed Control Indicator Scenarios

Three EACI-specific scenarios were developed. The goal was to induce a low-energy event by disguising an auto throttles failure, disconnect, or mode change with distractions or during a high workload flight condition such as landing in low visibility. The runs ended with either crew recognition of auto throttles not being engaged or recovery from stick-shaker with SRG flight director and thrust management guidance.

<u>Approach Scenario</u>

This scenario began with the aircraft flying over waypoint HOGGS along the Standard Terminal Arrival (STAR) CMRN4 towards a Category 3 Instrument Landing System (ILS) approach for Runway 4R at JFK in low visibility flight conditions. At 1,200 ft MSL, the auto throttles changed mode to IDLE at a rate of 0.5 deg/sec and then disengaged with no sound.

<u>Hold Scenario</u>

This scenario began with the aircraft flying over waypoint HOGGS along STAR CMRN4 towards an RNAV RNP approach to Runway 13L at JFK. At run startup, a data link message “SLOW TO 250 KNOTS, CROSS CAMRN AT 11,000 FEET” was issued. Passing through 12,000 ft MSL, a datalink message to “HOLD AT CAMRN AS PUBLISHED” was issued. At CAMRN, a distractor datalink message was issued for an unexpected approach type change to ILS 13L, causing the crew to go head-down to reprogram the FMS. As the aircraft descended in the hold through 11,500 ft MSL, the auto throttles changed mode to IDLE at a rate of 0.5 deg/sec and then disengaged with no sound.

Cruise Scenario

The scenario started with a departure on Runway 19 at DCA. At waypoint ACY, the auto throttles changed mode to IDLE at a rate of 0.5 deg/sec and when at 0.2 throttle lever position, the auto throttles disengaged. Simultaneously as the auto throttles went to IDLE, a distraction datalink message "TRAFFIC TWELVE O'CLOCK, TEN MILES OPPOSITE DIRECTION 19,000 FEET" was issued causing the crew to go head-up and eyes out of the flight deck to look for, identify, and assess the threat of the oncoming traffic.

3. Stall Recovery Guidance Scenarios

Training runs as well as three SRG-designed scenarios were employed to evaluate SRG. The training runs were used to measure: 1) a pilot's recovery from a full-stall condition using conventional 787-like displays (i.e., baseline condition) and using SRG-enabled displays; and, 2) a pilot's recovery from stick shaker (a near-stall condition) using SRG enabled displays. The training runs were described to the pilots "as an exercise to gather data in an operationally relevant environment for tuning of the SRG algorithms post-test." While this statement was true, it was also used to conceal from the pilots that "surprise stall" scenarios were specifically designed to evaluate the SRG concept during data collection. Cruise, departure, and approach scenarios were designed with the goal of having pilots unknowingly encounter a low-energy condition (approach to stall event indicated by stick shaker activation) from which they would recover with either SRG enabled or not enabled. Surprise stalls for an alert, well-trained crew are difficult to achieve especially when any normal deceleration of the aircraft occurs over a period of minutes. Simple distractions will not suffice. Only certain automation modes will permit aircraft stall, further complicating the scenarios.

Training Runs

After initializing the aircraft in a straight wings-level flight condition at 19,000 ft and 200 Knots Indicated Air Speed (KIAS) in a clean configuration, the pilots were instructed to pull the throttles back while maintaining altitude and heading. By doing so, they let the nose rise while slowing down, up to a pitch attitude of 15 deg (first runs) or 25 deg (later runs). This later pitch attitude angle was reached well beyond stick shaker into a full stall. Once this pitch target was reached, a stall recovery maneuver was executed to recover. The First Officer and Captain (in that order) did a training run without guidance first, followed by one with guidance for each one of them. The baseline condition (PFD without SRG guidance) was purposefully chosen to be flown before the guidance runs so that a pilot's trained stall recovery techniques were not influenced by the SRG pitch, roll, and thrust commands.

Cruise Scenario

This scenario was initialized enroute between Washington DC and Boston, at 25,000 ft MSL and 260 KIAS in a clean configuration. Two minutes into the run, ATC instructed an altitude change to FL350. Five seconds after the ATC instruction, the airspeed information on the PFD froze and when the vertical speed exceeded 1,500 ft/min, a drag increase was triggered. 55 seconds after the PFD airspeed information froze, an antiskid failure distracted the crew while the aircraft was bleeding airspeed, although not visible on the frozen airspeed tape of the PFD. When the stick shaker activated (which was driven by the stall warning angle of attack), the stall recovery guidance was engaged and the speed information became active again.

Departure Scenario

In this scenario, the aircraft was initialized holding short of Runway 31L at JFK airport. ATC instructed the crew to line up and wait, which was shortly followed by takeoff clearance. During the initial climb, additional drag was triggered at 1,000 ft above ground level. At the same time, an ATC voice call issued a traffic warning with the purpose of distracting the flight crew while the plane was losing speed. When the stick shaker activated, the stall recovery guidance was triggered and the flight crew performed a stall recovery maneuver.

Approach Scenario

This scenario began with the aircraft flying over waypoint HOGGS on the CMRN4 STAR towards a Category 3 ILS approach on Runway 4R of JFK. At 2 nm to waypoint EBBEE, voice messages on the tower frequency about an emergency at the airport was used to distract them. In the turn above EBBEE, the auto throttles retarded to IDLE and additional drag was introduced, which slowed the aircraft to a stall condition. At stick shaker, the crew initiated a stall recovery maneuver.

D. Test Variables and Procedures

The facility and environment were set up to fly the defined scenarios within the John F. Kennedy International Airport (KJFK) terminal airspace for arrivals, and at the Ronald Reagan Washington National (KDCA) Airport for departures. Available independent variables were weather conditions (5); airport configuration (2); traffic conditions

(3); flight path/procedure (9); and off-nominal conditions (8). Each scenario had unique settings for each of these variables to cover the selected span of test conditions for TRL assessment.

For arrivals, flights began just prior to the top-of-descent and on the published STAR procedure. A mix of ILS and RNAV approaches were flown to JFK. Prior to the start of each flight, the crew conducted an approach briefing, reviewed weather conditions at JFK, assigned Pilot Flying (PF) and Pilot Monitoring (PM) roles, and completed the descent ECL. During flight, pilots were asked to follow their airline's standard operating procedures as if on a revenue flight. Pilots were instructed to go-around if unstable or if they felt unsafe conditions existed and to continue flights until the researcher called "end of run."

For departures, flights began while holding short of the active (departure) runway and awaiting takeoff clearance. While holding short, all relevant checklists were accomplished as well as briefing NOTAMs and weather conditions. Takeoff and climb-out followed the appropriate published standard instrument departure (SID) procedure.

Both quantitative and qualitative data was recorded for AIME-2 to support post-hoc analysis. Data types included:

- Interview and questionnaire responses (both pilots, post-run and end-of-day)
- Head-/eye-tracking data (both pilots)
- Flight data (e.g., aircraft state, pilot button presses)
- Cockpit audio and video recordings
- Researcher and pilot observations

After each run, a short questionnaire was completed by each crewmember using a tablet computer. The questionnaire included the NASA Task Load Index (TLX) workload metric [25]; the Situation Awareness Rating Technique (SART) [26]; a Likert-type query on scenario complexity; pilot ranking of top three decision-making factors for maintaining safety of flight; and pilot selection of indicators used for energy state awareness, autoflight system state awareness, and system state awareness. For flights involving one of the technologies under evaluation, this questionnaire included the ten usability statements taken from the System Usability Scale [27] and a Likert-type query regarding acceptability.

E. Subjects

26 commercial airline pilots (13 crews), representing three U.S. airlines, participated in the experiment. All subjects (22 male, 4 female) were Airline Transport Pilot rated and currently qualified in wide body commercial aircraft. Captains had an average of 24,288 commercial flight hours with eight having an average of 16.9 years military flight experience. First Officers had an average of 11,433 commercial flight hours with nine having an average of 15.3 years military flight experience. Captains and First Officers were paired by airline to ensure crew coordination and cohesion with regard to airline standard operating procedures.

F. Summary of Completed Tests

The 13 crews flew 26 new technology/simulator familiarization runs, 78 SRG training runs for post-test algorithm development [19, 20], and 102 data collection runs.

A between-subjects experiment design was used for the three ASA technologies tested. Eight data collection runs were planned for each crew: 2 system failure events, 3 EACI-specific scenarios, and 3 SRG-specific scenarios.

Odd-numbered crews were assigned the pitot-static failure with the baseline display condition and the hydraulic failure with the eSyn/sECL display condition. Even-numbered crews were assigned the pitot-static failure with the eSyn/sECL display condition and the hydraulic failure with the baseline display condition.

Each crew flew two EACI scenarios with the EACI display condition and the remaining EACI scenario with the baseline condition. The baseline and EACI display conditions were equally distributed across the three EACI scenarios.

Similarly, each crew was assigned two SRG scenarios with the SRG display condition and the remaining SRG scenario with the baseline condition. The baseline and SRG display conditions were equally distributed across the three SRG scenarios. Algorithm type was held constant for each crew. Odd-numbered crews used the model-predictive control approach algorithm and even-numbered crews used the energy-based approach algorithm.

Table I provides a summary of the 102 data collection flights completed by the 13 crews across the various scenarios and display conditions (baseline, eSyn/sECL, EACI, SRG).

Table I. Number of Runs for Each Combination of Scenario and Display Condition

Scenarios	*Baseline Display*	*Enhanced Displays (eSyn/sECL, EACI on PFD, or SRG on PFD)*
Pitot-Static Failure	6	7
Hydraulic Failure	6	6
EACI-1 (approach)	4	9
EACI-2 (hold)	6	7
EACI-3 (cruise)	3	10
SRG-1 (approach)	3	9
SRG-2 (departure)	4	9
SRG-3 (in holding pattern)	6	7

III.Selected Findings

Selected analyses of usability ratings, acceptability ratings, situation awareness ratings, and workload ratings as well as pilot feedback regarding the three technologies are presented here. The eSyn/sECL results presented here are for the left hydraulic system failure runs only. The EACI analyses are collapsed across the scenarios and the SRG analyses are collapsed across scenarios and algorithm type.

Selected quantitative results presented here include the following. One measure of eSyn/sECL performance was characterized by the time to complete the Hydraulic Systems Pressure (Left Only) non-normal checklist. EACI performance was characterized by time elapsed and speed loss before crew recognition that automation was not directly controlling to a target airspeed. Additionally, analyses of two runs are discussed where a crew using the EACI concept failed to recognize the decreasing airspeed and had a “surprise” stick shaker event. Full stall was avoided using the SRG technology in these cases.

A. Usability and Acceptability

The System Usability Scale (SUS) [27] was used to gauge how pilots assessed the usability of the three ASA technologies tested. Using the method described in [27], SUS scores were calculated based on ten sub-scores and fell in a range from 0 to 100, but these are not percentile ranks. SUS scores can be associated with specific letter grades and adjective ratings [28]. A SUS score between 63 and 80.3 is considered a “good” design and a score above 80.3 is considered an “excellent” design.

Acceptability of the ASA technology being used was self-assessed after each run using a seven point Likert scale with a rating of 1 being “very unacceptable,” a rating of 7 being “very acceptable,” and a rating of 4 being “average.”

In Fig. 6, mean PF and PM System Usability Scale (SUS) scores and technology acceptability ratings are provided for the eSyn/sECL (hydraulic failure runs only), EACI, and SRG technologies. SUS scores, means with ± 2 standard errors (SE), are shown as multi-color bar graphs (referencing left y-axis scale) and acceptability scores are shown as red circles with mean values listed adjacent to them (referencing right y-axis scale).

Pilots, in general, found the design of the eSyn/sECL to be a good pilot interface tool (Fig. 6, mean SUS score of 87 for the PF and 80 for the PM) for dealing with non-normal flight situations like a hydraulic failure. Additionally, self-reported ratings revealed both the PF (mean value 6.0) and PM (mean value of 5.8) found the eSyn/sECL to be highly acceptable for dealing with the hydraulic failure event.

Both the PF and PM considered the EACI and SRG enhanced PFDs to be excellent designs (Fig. 6, mean SUS ratings all above 80.3). Additionally, self-reported mean ratings of 5.8 or better revealed that pilots, regardless of flight role, found the EACI and SRG display conditions to be highly acceptable.

Self-reported post-run ratings indicated that pilots judged the pilot interfaces to be usable and acceptable for the eSyn/sECL, EACI, and SRG concepts. Interestingly, although the eSyn/sECL and SRG were designed to primarily support the PM and PF respectively, this data shows no significant difference in how they were judged for usability and acceptability.

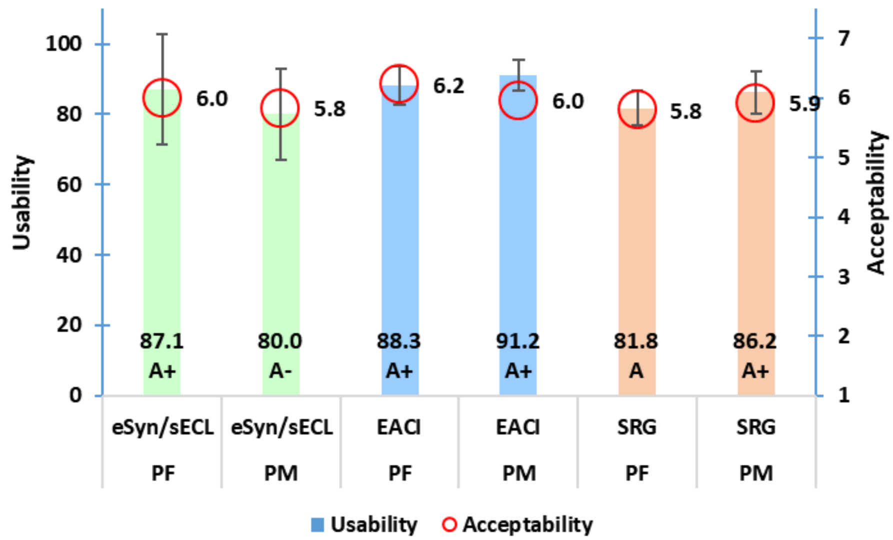

Fig. 6 PF and PM usability and acceptability ratings for AIME-2 ASA technologies

B. Situation Awareness and Workload

The Situation Awareness Rating Tool (SART) [26] is a multi-dimensional rating technique using the constructs of: 1) demand on attentional resources; 2) supply of attentional resources; and, 3) understanding. From these components, the SART rating is "understanding" reduced by the difference of "demand" minus "supply" (i.e., SART = {(understanding) – (demand – supply)}). The SART technique provides relative situation awareness pilot ratings among the display conditions.

The NASA Task Load Index (TLX) method captures a subjective rating (0 - "Low" to 100 - "High") of perceived task load. There are six subscales of workload represented in the NASA TLX: mental demand, physical demand, temporal demand, performance, effort, and frustration level [25]. Overall workload was calculated as the unweighted average of the ratings of the six subscales for both the PF and PM and was used to examine task load variation between the baseline condition and ASA technologies tested.

Based on the questionnaire responses using SART and TLX, the resulting PF and PM situation awareness and workload ratings are shown in Fig. 7-Fig. 9 for the eSyn/sECL, EACI, and SRG technologies, respectively. The baseline condition (i.e., the three ASA technologies were not enabled) is indicated as "BL" in Fig. 7-Fig. 9.

For the hydraulic failure runs with eSyn/sECL, pilots rated their overall workload (Fig. 7, right side) as being moderate as reflected in the PF (mean rating of 46) and PM (mean rating of 58) TLX ratings. There were no significant PF or PM workload differences between the baseline and eSyn/sECL display conditions. Both the PF and PM reported situation awareness gains (Fig. 7, left side) with the eSyn/sECL compared to the baseline for dealing with a hydraulics failure.

For the EACI runs, pilots rated their overall workload (Fig. 8, right side) as being moderate as reflected in the PF (mean rating of 42) and PM (mean rating of 38) TLX ratings. There were no significant PF or PM workload differences between the baseline and EACI display conditions; however, EACI did show trends for slightly lower reported workload levels than the baseline condition. Both the PF and PM reported situation awareness gains (Fig. 8, left side) with the EACI-enhanced PFD compared to the baseline PFD.

For the SRG runs, pilots rated their overall workload (Fig. 9, right side) as being moderate as reflected in the PF (mean rating of 55) and PM (mean rating of 47) TLX ratings. There were no significant PF or PM workload differences between the baseline and SRG display conditions. Both the PF and PM reported slightly lower SART ratings (Fig. 9, left side) for the SRG-enhanced PFD compared to baseline PFD.

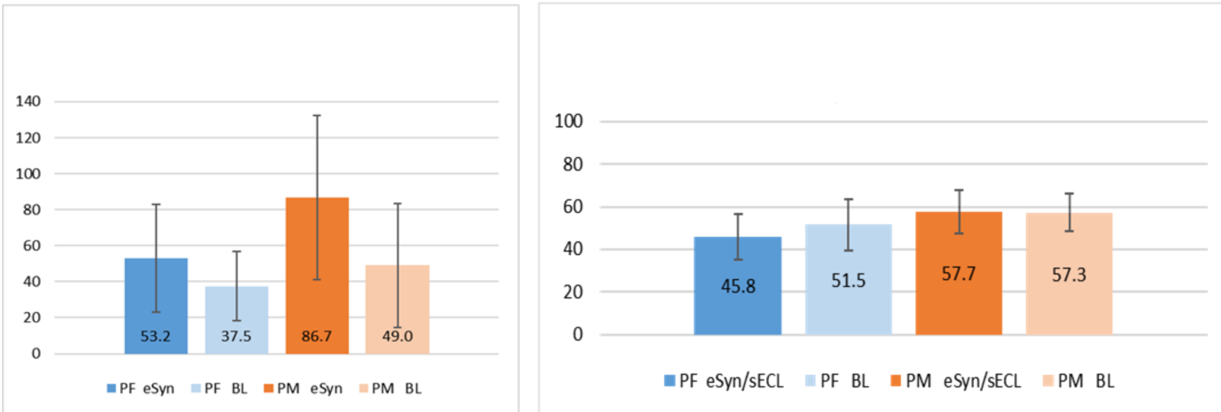

Fig. 7 PF and PM SART (left) and TLX ratings (right) for hydraulic failure scenario (mean, ±2 standard errors)

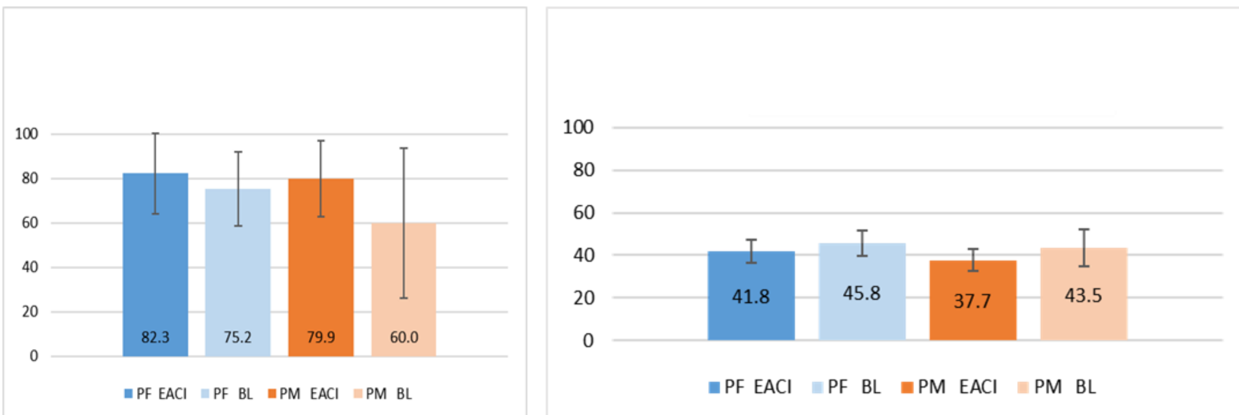

Fig. 8 PF and PM SART ratings (left) and TLX ratings (right) for EACI scenarios (mean, ±2 standard errors)

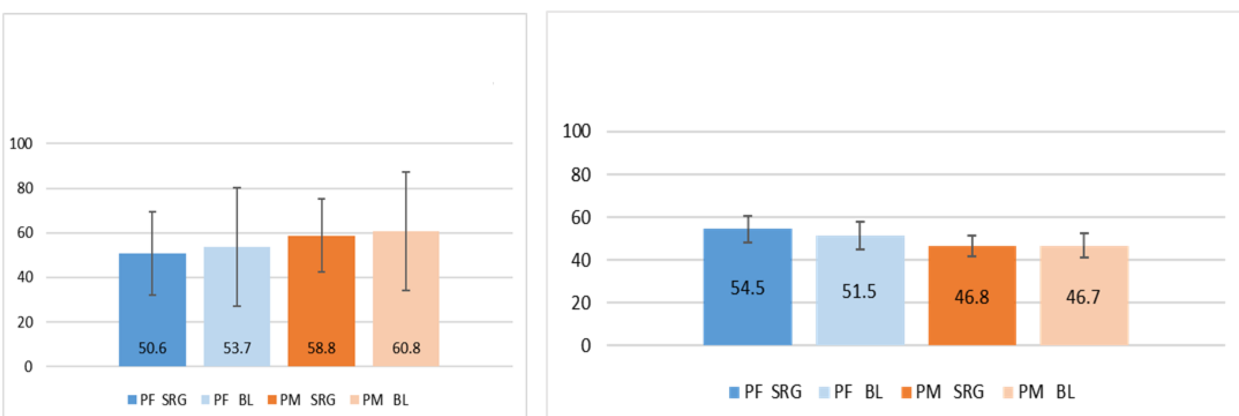

Fig. 9 PF and PM SART ratings (left) and TLX ratings (right) for SRG scenarios (mean, ±2 standard errors)

Post-run ratings indicated PF and PM situation awareness gains without any reported workload penalties for the eSyn/sECL and EACI concepts as compared to the baseline condition. Self-reported post-run ratings indicated a slight decrease in situation awareness and marginal increase in workload for both PF and PM when using the SRG compared to the baseline condition; however, these differences were not statistically significant.

C. Pilot Comments

Each crew participated in a post-test interview and debriefing where they also completed a short questionnaire asking them to provide comments regarding their overall satisfaction with the information provided by the new technology concepts. Discussion and selected observations from these are provided below for each.

1. Enhanced Synoptics and Simplified Electronic Checklist

In general, pilots favored using the enhanced flight control and hydraulic synoptic pages in concert with a simplified electronic checklist for the left hydraulic failure employed in this study. Pilot comments supported the SUS ratings that found the technology to be a good system design for understanding the effects of such a failure on the aircraft and aircraft systems. Example written comments include:

> *"This very much improves understanding of the state of the aircraft system, both the malfunction and how it affects aircraft control. Having it graphically displayed and with short phrases ("do not autoland") right on the display in amber made it easy to understand what we were dealing with and how to proceed. Even as PF, I could glance over and see it."*

> *"Excellent improvement. Really like the landing distance calculations."*

> *"Overall, satisfied with the concept and the display and I did feel it increased my SA concerning the energy state."*

> *"It is a nice enhancement with ever increasing system complexity. Quick access to important flight information."*

> *"Better than standard. The pictorials are easier to understand then written word. Especially performance numbers."*

> *"Highly beneficial in visualizing and understanding certain system malfunctions!"*

> *"This is great and should be implemented ASAP. Everything that you've lost is condensed on one page with landing distances, so you don't have to search through QRHs to find numbers."*

2. Enhanced Airspeed Control Indicator

In general, pilots commented that the EACI concept was a useful technology for indicating to them when target airspeed was not being actively controlled by the automation; however, many pilots recommended additional notification strategies (e.g., aural, visual) should be explored to augment the EACI display elements currently employed. Example written comments include:

> *"Really liked the white XX's displayed on the PFD speed tape. It was clear, and alerted me quickly to uncontrolled airspeed."*

> *"The indications that the EACI provide are necessary info for situation awareness and would increase flight safety. Add an aural would make the system even better. I am very satisfied with this system."*

> *"The X's appearing do give you a visual clue that the speed mode has failed. Audible indication should be added in case the PF is not looking at PFD."*

> *"Helpful enhancement, but I didn't notice it right away. Perhaps an aural tone or flashing indicator for a few seconds would be a good cue to notice it. Overall a good indication."*

> *"Great visual indication for failed auto throttles. I would add XXs to the FMA as well"*

> *"I think this is an improvement that will help crews to identify an auto-throttle or related problem affecting airspeed - however it would be improved by an audible warning - I noticed that each new datalink message had a chime that caught my attention but this issue had only a visual alert and no aural alert."*

3. Stall Recovery Guidance

In general, pilot comments supported the SUS ratings that indicated that the stall recovery guidance technology was an excellent system design. Example written comments include:

"Excellent System. Great guidance. Long overdue!"

"To me, very similar to the windshear recovery system most are used to. Throttle indication is outstanding for recovery and recovery of the recovery."

"I like the SRG as a means to focus on the critical part of stall recovery with a simple indicator akin to the PFD indication during a windshear event. I am very satisfied with this system."

"Fantastic. This is visual so most (pilots) people can quickly assess the situation and correct for it in minimal time and altitude loss."

Pilots also provided suggested improvements to the SRG display presentation.

"Very nice overall; make throttle band bigger; gauge or display symbology to show required vs requested thrust."

"Mostly satisfied, but had problems with bringing the power back; when the indication was to bring the throttles back showed a speed close to the VREF Speed on the PFD speed tape (top of yellow band)."

"Good system - would like the mode indicators to blank when aircraft is back to normal flying mode."

"Not entirely sure I like the throttle cues. The indication for power reduction does not tell me what state the system is recovering back to, or when to exit the mode."

In addition to display suggestions, pilots provided feedback on system performance and integration with existing flight deck warning systems.

"Great tool to assist recovery. Please adjust post recovery speed to clean maneuvering vice 180. With the stress of the recovery, you get focused on altitude first and climbing away at 180 knots can be undesirable."

"This is an improvement in that it combines pitch, roll, and thrust guidance that is combined on the PFD. The only question in my mind was how well integrated it was with the EGPWS ('Pull up')."

D. Quantitative Performance Indicators

Several quantitative measures can help to indicate to what extent a new technology can affect performance and its usefulness for improving ASA. Selected examples are given below for the three technologies evaluated in AIME-2.

1. eSyn/sECL Checklist Completion Time

Time to execute relevant checklist(s) can be an indicator for effectively handling failures and understanding their effects. The notched boxplots in Fig. 10 indicate that the median checklist completion times were significantly different (0.05 significance level) when using eSyn/sECL (mean time of 537 sec or 8 min 57 sec) versus when only using the baseline checklist (mean time of 774 sec or 12 min 54 sec) for the left hydraulic systems failure. On average, crews were able to complete the non-normal checklist almost four minutes sooner using the eSyn/sECL technology as compared to using the baseline Synoptics/ECL. This is ~30% time reduction. The maximum time savings across all runs was 42% (6 min and 10 sec).

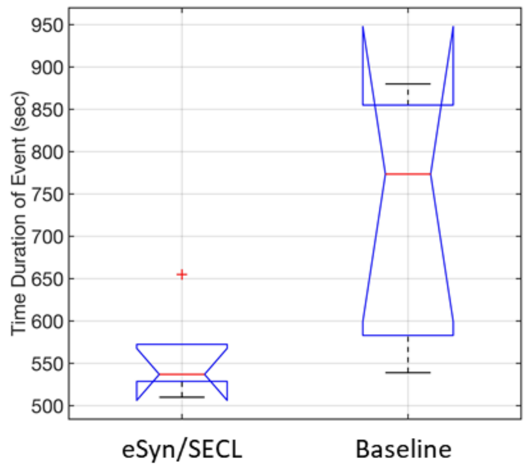

Fig. 10 Checklist completion time for hydraulics system failure using eSyn/sECL vs baseline

2. EACI Performance for the A/T Disengagement during Holding Pattern Scenario

Table II provides the time elapsed and airspeed loss from when the auto throttles disengaged until the crew recognized the situation either by advancing the throttles or stick shaker activation for the EACI holding pattern scenario runs. All crews noticed the auto throttle state change in the baseline condition (mean recognition time of 32.1 seconds with 24 knots of airspeed lost) before stick shaker, but only 5 of the 7 crews recognized the state change with the EACI-enhanced PFD (mean recognition time of 16.6 seconds with 9 knots of airspeed lost) before stick shaker. As shown in Fig. 11, two of the crews flying with EACI experienced a stick shaker event before noticing the loss of airspeed control by the auto throttles, but did recover using the SRG.

Current PFDs have a lot of information on them and information is often looked at but not processed. Post-run subjective ratings indicated that the relatively simple EACI interface (white XX's near three key PFD airspeed location elements) was usable and acceptable but pilot comments revealed that additional notification strategies for when automation was not actively controlling to a target airspeed may be desirable. The two unplanned approach to stall events while flying with the EACI-enhanced PFD appear to support these pilot comments. The limited set of data gathered suggests that when the EACI information was processed crews recognized the decaying airspeed approximately two times faster than when using the baseline condition.

Table II. Time to recognize auto throttles disengagement and airspeed loss

Display Condition	*Recognition time (sec)*	*Speed Loss (kts)*
Baseline	25.6	14.2
Baseline	77.4	65.8
Baseline	22.1	25.6
Baseline	16.4	11.6
Baseline	24.4	8.5
Baseline	26.8	18.3
EACI*	72.4	68.9
EACI*	101.6	86.1
EACI	47.3	32.5
EACI	8.1	0.7
EACI	1.6	-0.3
EACI	9.2	1.6
EACI	16.8	10.8

* Stick shaker event occurred before crew recognition of inactive auto throttles.

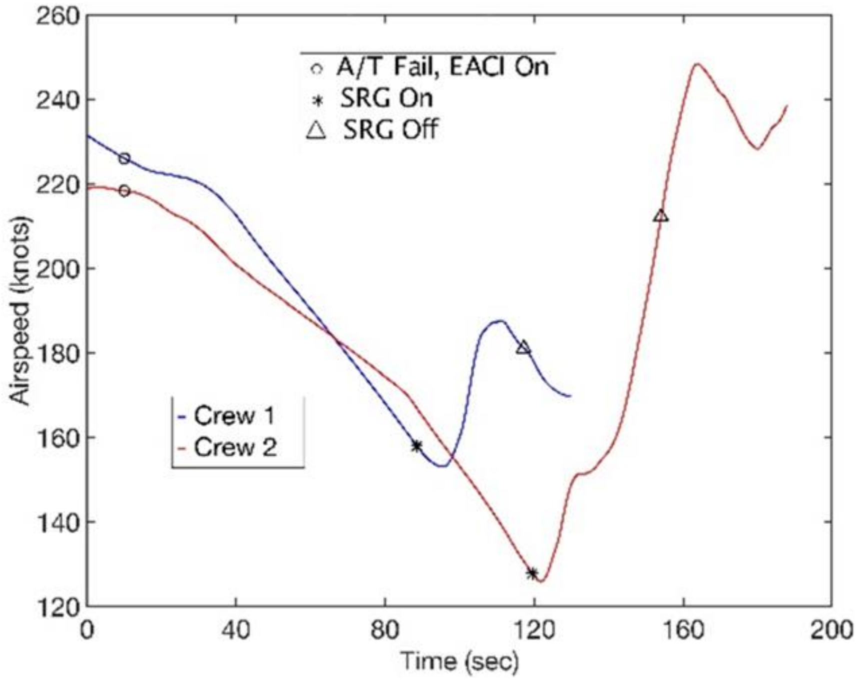

Fig. 11 Airspeed recovery from stick shaker event

3. SRG Performance for Unplanned EACI Stick Shaker Events

SRG criteria [20] for successfully recovering from approach-to-stall or stall include:

- not exceeding the load factor n_z,
- no secondary full stalls (angle of attack not significantly more than stall warning angle of attack)
- avoiding excess altitude loss during recovery

For the two unexpected/unplanned approach to stall warning events experienced by crews flying with the EACI, SRG was effective in providing a pitch reference for the crews to quickly break the stall angle of attack and efficiently recover to a safe airspeed without excessive load factors or secondary full stalls. They were able to successfully recover, without excessive altitude loss, even with the limited SRG training received prior to data collection.

IV. Discussion

With respect to dealing with system failures and understanding their effects, data collected in AIME-2 and previously in AIME-1 [13] shows that the time to complete the appropriate checklist can be significantly reduced using the eSyn/sECL concept. In this study, checklist execution time was reduced by a third. Further, with the graphical visual available to both crewmembers, the PF does not have to comprehend what the PM is reading and the information remains available on the synoptic page throughout the rest of the flight and can be reviewed at any time. There's no quantitative performance measure for this capability. Although no procedural steps were removed, the checklists were simplified by removing notes that described in textual form which items were failed or degraded. Instead, the items were depicted graphically on the synoptic page, and a note on the checklist directs to the crew to 'Refer to the Synoptic for Inoperative Items and Effects'. This alone can result in significant reduction in the length of the checklists and has a number of added benefits, even when/if the entire checklist is only reduced by one page. Key checklist steps are often elevated to the first page of the checklist when notes are reduced so comprehension of the failure and the effects are known quickly. Although hard to replicate in a simulator, checklists are interrupted in many ways in actual flight conditions. Communications from ATC and airline operations center personnel trying to help are hard to ignore. When items can be completed more quickly, the procedure will often be significantly safer. Although PM usability scores only reflect a good design, PF scores reflected an excellent design. This is not surprising since the PM role of completing the checklist was not significantly simplified but providing simple to understand failure effects supported the role of the PF and the information was persistent. The equivalent of graphical notes allowed the PF to evaluate failure effects immediately when the eSYN page was displayed. Although the current study was terminated after completion of the checklist, if taken to a landing, the approach/landing checklists required a review of notes. The reduction of reviewing the notes for the landing phase would have supported the PM as well. The support provided by the eSyn/sECL to the PF was evident in the reduction of workload and increased situation awareness and understanding of current airplane conditions and limitations. Although only tested for a couple of failure types, it is believed this approach can be extended to other failure types and involve other existing (or new) synoptic pages.

Lack of energy awareness can be indirectly observed by delay in noticing or taking appropriate action following an A/T mode change or disconnect. This is also evident by the total airspeed loss during this time frame. The relatively simple EACI design of augmenting the PFD with white XX's near three key airspeed information elements to indicate whether automation was actively following the target airspeed was judged to be acceptable and useable by the pilots without any increases in workload. Subjective ratings are one part of measuring acceptability and usability of an ASA technology in an operationally relevant environment, but objective measures of the ASA effectiveness/usefulness must also be considered. The limited data reported in this paper shows a mixed message with respect to the EACI effectiveness in indicating to the crew when automation was no longer actively controlling (e.g., inactive auto throttles) to the target airspeed. When recognized and processed, the simple display alerting strategy of the EACI was highly effective as detection time to notice decaying airspeed was two times quicker than when using the baseline. The average airspeed loss was nine knots with EACI and 24 knots with the baseline before crew recognition of the failure. Airspeed loss of 10 knots or greater is considered operationally relevant. The magnitude of the airspeed loss in the baseline condition is not surprising since crews did not notice the FMA change and would only notice the auto throttle state change after airspeed diverged from the selected airspeed bug.

However, there were two flights with crews using the EACI-enhanced PFD that approached stall conditions. It is recommended (and supported by pilot comments) that additional notification strategies (e.g., aural, visual) should be explored to augment the EACI display elements currently employed.

The focus of the AIME-2 experiment for SRG was on pilot evaluation under operationally complex scenarios. This scenario complexity went significantly beyond the previous study [29, 30] where the experiment objective was to evaluate recovery guidance algorithm performance on full stall recovery from at or beyond the stall reference angle of attack. The full stall recovery scenarios of this previous study were, however, simpler and more consistently repeatable. To extend the test evaluation envelope for SRG, the AIME-2 experiment focused on recovery from stall warning (at the stall warning angle of attack, which is typically less than the stall reference angle of attack by several degrees), where pilots would normally initiate a recovery. Overall, the pilot comments favored the use and benefit of the SRG system, while noting a few areas for improvement. This was quantitatively established on the SUS (Fig. 6), with no significant effect on workload (Fig. 9). Even with limited training, the SRG was effectively used by the crews to safely recover from unplanned (or surprise) approach to stall conditions without excessive load factors or altitude loss. Additionally, AIME-2 pilot ratings of the technology, flown with significantly increased scenario complexity, were consistent with those found in the previous SRG study which established algorithm performance on full stall recovery [29, 30]. Overall, these observations support raising the TRL of the SRG technology.

V. Conclusions

All three technologies evaluated were deemed usable and acceptable by the participating flight crews. Further, evaluations were conducted in a relevant environment and spanned a set of complex situations such as encountered by crews in previous events. Together, this raises the TRL to five and achieves the CAST goal for relevant research outputs. Typically, the next step for this TRL is consideration for transfer and inclusion in future aircraft designs, or perhaps retrofit.

With respect to Synoptics, current non-normal procedures do not require reference to Synoptic pages, even in aircraft where these displays are part of the type certificate. If, as suggested in this paper, references to Synoptic pages replace more detailed prose on checklists, the Synoptic display may call for additional certification costs (i.e., the synoptic becomes essential equipment to complete a non-normal checklist). A risk-based approach for certification is now in place for Part 23 aircraft, perhaps the same approach can be applied to Part 121 aircraft for secondary information like Synoptic displays allowing reductions in certification costs for technology that has a clear and proven safety benefit. This may allow cost-effective retrofit of such technology.

Stall recovery guidance provides a number of benefits. First, it provides a positive indication for a pitch attitude that will break a stall. Second, it provides a recovery maneuver that reduces the risk of secondary stalls and excessive load factors. Although not studied here, the guidance will provide an alert indication if the current stab trim setting will inhibit proper recovery. Stall and approach to stall is currently a training focus in the airlines. A number of PM crew members had completed the training and verbalized decreasing/increasing airspeed, altitude increasing, etc. during the runs with and without SRG. SRG did not interfere with these learned procedures and appeared to support the PF during the recovery.

Visual only enhancements may not be completely effective for energy state awareness. Current PFDs are necessarily cluttered and information is often looked at but not processed. Although pilot ratings indicated situation awareness increases with EACI this was not the case for some crews where the enhancement was missed for nearly two minutes as the aircraft encountered stick shaker.

Several measures and data collected during AIME-2 are not reported here. A separate paper on eye-tracking results is planned; and there will be reporting of findings not specific to the new technologies being evaluated, but rather the general and long-term trend to increasingly complex flight deck systems and information automation.

Acknowledgments

The authors would like to thank Captain Daniel Kiggins of American Airlines for his help with scenario development at LaRC and at the American Airlines Boeing 787 Training Simulator. Thanks also to the CMF/RFD simulation development team, led by Victoria Chung and Sean Kenney, for their creativity and dedication in implementing our software and hardware requirements for the execution of the AIME-2 experiment.

This work is part of ongoing collaborative research under an inter-agency agreement between NASA and the Federal Aviation Administration. Special thanks also to Sabreena Azam, Ashley Awward, and John Stuber of the FAA for their support and guidance along the way.

References

[1] Commercial Aviation Safety Team; "Airplane State Awareness Joint Safety Analysis Team Final Report," Skybrary Bookshelf, Jun 2014.

[2] Commercial Aviation Safety Team; "Airplane State Awareness Joint Safety Implementation Team Final Report," Skybrary Bookshelf, Jun 2014.

[3] Commercial Aviation Safety Team; "Safety Enhancement SE 207 ASA Research – Attitude and Energy State Awareness Technologies," Skybrary Bookshelf, Jun 2014.

[4] Commercial Aviation Safety Team; "Safety Enhancement SE 208 ASA Research – Airplane Systems Awareness," Skybrary Bookshelf , Jun 2014.

[5] Young, S.; Daniels, T; Evans, E.; Shish, K.; Schuet, S.; Etherington, T.; Uijt de Haag, M.; and Kiggins, D.; "Evaluating Technologies for Improved Airplane State Awareness and Prediction" AIAA Science and Technology Forum and Exposition (SciTech), San Diego, CA, January 4-8, 2016.

[6] Young, S.; Daniels, T.; Evans, E.; Dill, E.; Uijt de Haag, M.; and Etherington, T.; "Flight Simulation Study of Airplane State Awareness and Prediction Technologies," Proceedings of the AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, September 25-29, 2016.

[7] Shish, K.; Kaneshige, J.; Acosta, D.; Schuet, S.; Lombaerts, T.; Martin, L.; and Madavan, A.; "Trajectory Prediction and Alerting for Aircraft Mode and Energy State Awareness," paper no. AIAA 2015-1113, AIAA SciTech 2015, Kissimmee, FL, Jan 2015.

[8] Lombaerts, T.; Schuet, S.; Acosta, D.; Kaneshige, J.; and Shish, K.; "Piloted Simulator Evaluation of Maneuvering Envelope Information for Flight Crew Awareness," paper no. AIAA 2015-1546, AIAA SciTech 2015, Kissimmee, FL, Jan 2015.

[9] Uijt de Haag, M.; et. al.; "Energy State Prediction Methods for Airplane State Awareness," Proceedings of the AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, September 25-29, 2016.

[10] Dill, E.; Young, S.; Daniels, T.; and Evans, E.; "Analysis of Eye-Tracking Data during Conditions Conducive to Loss of Airplane State Awareness," Proceedings of AIAA AVIATION Forum, Denver, CO, June 5-9, 2017.

[11] Neville, R. and Dey, M., "Innovative 787 Flight Deck Designed for Efficiency, Comfort, and Commonality," Boeing Commercial Aero Magazine, Quarter 01.12, 2012.

[12] Etherington, T., Kramer, L. J., Le Vie, L. R., Last, M. C., Kennedy, K. D., Bailey, R. E., and Houston, V. E., "Impact of Advanced Synoptics and Simplified Checklists during Aircraft Systems Failures," in Digital Avionics Systems Conference (DASC), 2018 IEEE/AIAA 37th, 2018, pp.670-678.

[13] Young, S. D., Uijt De Haag, M., Daniels, T., Evans, E., Shish, K. H., Schuet, S., Etherington, T., and Kiggins, D., "Evaluating Technologies for Improved Airplane State Awareness and Prediction," in AIAA Infotech Aerospace. AIAA SciTech, 2016, no. AIAA 2016-2043, p. 2043.

[14] Young, S. D., "Airplane State Awareness and Prediction Technologies – Research Overview (2014-2018)," in AIAA Science and Technology Forum and Exposition SciTech, 2019

[15] Lombaerts, T., Schuet, S., Kaneshige, J., Shish, K., and Stepanyan, V., "Stall Recovery Guidance Using an Energy Based Algorithm," AIAA SciTech Conference, No. AIAA-2017-1021, American Institute for Aeronautics and Astronautics, 2017.

[16] Schuet, S., Lombaerts, T., Kaneshige, J., Shish, K., and Stepanyan, V., "Stall Recovery Guidance Using Fast Model Predictive Control," AIAA SciTech Conference, No. AIAA-2017-1513, American Institute for Aeronautics and Astronautics, 2017.

[17] Sherry, L. and Mauro, R., "Design of Cockpit Displays to Explicitly Support Flight Crew Intervention Tasks," in Digital Avionics Systems Conference (DASC), 2014 IEEE/AIAA 33rd, 2014, pp. 2B5.1–2B5.13.

[18] Daniels, T. S., Ferguson, C. M., Dangtran, E. T., Korovin, R. M., Kramer, L. J., Evans, E. T., Santiago-Espada, Y., Kiggins, D., Etherington, T. J., and Barnes, J. R., "Regarding Pilot Usage of Display Technologies for Improving Awareness of Aircraft System States," in IEEE Aerospace Conference, Yellowstone Conference Center, Big Sky, Montana, 2019.

[19] Schuet, S., Lombaerts, T., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., Robinson, P., Etherington, T., Kramer, L. J., Evans, E. T., Daniels, T. S., and Young, S. D., "Piloted Simulation Study Findings on Stall Recovery Guidance," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[20] Lombaerts, T., Schuet, S., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., Robinson, P., Kramer, L., Etherington, T., Daniels, T., Evans, E. T.,and Young, S., "Design and Piloted Simulator Evaluation Results of Model Independent Stall Recovery Guidance," in AIAA Science and Technology Forum and Exposition SciTech, 2019.

[21] Young, S., Daniels, T., Evans, E., and Uijt de Haag, M., "Understanding Crew Decision Making in the Presence of Complexity – A Flight Simulation Experiment," paper no. AIAA 2013-4894, AIAA InfoTech@Aerospace Conference, Boston, MA, Aug 2013.

[22] NTSB, "Descent Below Visual Glidepath and Impact With Seawall – Asiana Airlines Flight 214," NTSB Accident Report AAR1401, Jun 2014.

[23] BEA, "Air France 447 Accident Report," French Bureau of Enquiry and Analysis, Jul 2012.

[24] Finland Accident Investigation Board, "Serious Inchident at Helsinki-Vantaa Airport due to Tread Separation and Hydraulic Failure, 22 June 2009", Investigation Report C2/2009L, April 2010.

[25] Hart, S.G. and Staveland, L.E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," In P.A. Hancock & N. Meshkati (Eds.), Human Mental Workload. Amsterdam: North-Holland, 1988, pp. 139-183.

[26] AIAA, “Guide to Human Performance Measurement,” Washington D.C.: American Institute of Aeronautics and Astronautics, 1993.

[27] Brooke, J.; “SUS – A Quick and Dirty Usability Scale,” in P. Jordan, B. Thomas, B. Weerdmeester, A. McClelland, Usability Evaluation in Industry, London: Taylor & Francis, 1996.

[28] Sauro, J., A Practical Guide to the System Usability Scale, Denver: Measuring Usability, 2011.

[29] Schuet, S., Lombaerts, L., Stepanyan, V., Kaneshige, J., Shish, K., Robinson, R., and Hardy, G., “Vertical Motion Simulator Experiment on Stall Recovery Guidance,” NASA Technical Report NASA/TP–2017–219733, National Aeronautics and Space Administration, October 2017.

[30] Lombaerts, T., Schuet, S., Stepanyan, V., Kaneshige, J., Hardy, G., Shish, K., and Robinson, P., "Piloted Simulator Evaluation Results of Flight Physics Based Stall Recovery Guidance", 2018 AIAA Guidance, Navigation, and Control Conference, AIAA SciTech Forum, (AIAA 2018-0383).